SICE Annual Conference 2012
August 20-23, 2012, Akita University, Akita, Japan

High-Accuracy Image Matching Using Phase-Only Correlation and Its Application

Mamoru Miura, Shuji Sakai, Shoichiro Aoyama, Jumpei Ishii, Koichi Ito and Takafumi Aoki

Graduate School of Information Sciences, Tohoku University, Sendai, Japan
(Tel: +81-22-795-7169; E-mail: miura@aoki.ecei.tohoku.ac.jp)

Abstract: Image matching is the most important fundamental task in many fields, such as image sensing, video signal processing, computer vision, etc. Recently, the demand for high-performance image matching achieving sub-pixel registration is rapidly growing in many applications. We have developed an efficient image matching technique using Phase-Only Correlation (POC), which uses the phase components in Discrete Fourier Transforms (DFTs) of given images. In this paper, we briefly present fundamentals of our image matching technique and examples of the potential applications.

Keywords: image matching, 3D measurement, stereo vision, biometrics, phase-only correlation

1. INTRODUCTION

Image matching is of fundamental importance in many fields, such as image sensing, video signal processing, computer vision, etc. Recently, the demand of high-performance image matching achieving sub-pixel registration accuracy is rapidly growing in many applications. For example, narrow-baseline stereo-vision systems generally require efficient sub-pixel correspondence algorithms to achieve high accuracy in 3D reconstruction. Also, sub-pixel motion estimation is essential for super-resolution imaging, which is to reconstruct a high-resolution image from multiple low-resolution images.

Image registration techniques using phase information of discrete Fourier transform could address requirements of sub-pixel registration. In 1975, Kuglin published a pioneering research on phase correlation [1]. This technique was successfully applied to the estimation of image transformation parameters (i.e., translation, rotation and scaling) via Fourier-Mellin transform [2], and video motion estimation [3]. While some researches use the terms phase correlation and phase-only matched filtering, our research group uses the term Phase-Only Correlation (POC), for historical reasons. All these terms have been used almost in the same meaning. Also, the works on MACE (Minimum Average Correlation Energy) filters and other correlation filters [4] are closely related to POC-based techniques.

Since 1990s, our research group has developed a novel technique of POC-based image matching for biometric authentication, achieving commercial application of a series of fingerprint verification systems [5]. The same technique has also been successfully applied to high-speed image recognition systems for industrial machine vision applications [5]. On the basis of these developments, we have recently proposed an efficient image correspondence algorithm using POC, which can find pairs of corresponding points between the given two images with sub-pixel accuracy [6]. The proposed algorithms allow us to apply the POC technique to a wide range of applications, such as image sensing, computer vision, industrial image recognition, biometrics, and waveform analysis [5-21]. Fig. 1 shows various applications of the POC technique. In this paper, we present fundamentals and potential applications of the POC technique.

2. PHASE-BASED IMAGE MATCHING

This section describes the fundamental definition of Phase-Only Correlation and an image matching technique using POC.

Consider two $N_1 \times N_2$ images, $f(n_1, n_2)$ and $g(n_1, n_2)$, where we assume that the index ranges are $n_1 = -M_1, \cdots, M_1$ $(M_1 > 0)$ and $n_2 = -M_2, \cdots, M_2$ $(M_2 > 0)$ for mathematical simplicity, and hence $N_1 = 2M_1+1$ and $N_2 = 2M_2+1$. Let $F(k_1, k_2)$ and $G(k_1, k_2)$ denote the 2D DFTs of the two images. $F(k_1, k_2)$ and $G(k_1, k_2)$ are given by

$$\begin{aligned} F(k_1,k_2) &= \sum_{n_1,n_2} f(n_1,n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_F(k_1,k_2) e^{j\theta_F(k_1,k_2)}, \qquad (1) \\ G(k_1,k_2) &= \sum_{n_1,n_2} g(n_1,n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_G(k_1,k_2) e^{j\theta_G(k_1,k_2)}, \qquad (2) \end{aligned}$$

where $k_1 = -M_1, \cdots, M_1$, $k_2 = -M_2, \cdots, M_2$, $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$, $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$ and $\sum_{n_1,n_2}$ denotes $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2}$. $A_F(k_1, k_2)$ and $A_G(k_1, k_2)$ are amplitude and $\theta_F(k_1, k_2)$ and $\theta_G(k_1, k_2)$ are phase. the normalized cross-power spectrum $R_{FG}(k_1, k_2)$ is given by

$$\begin{aligned} R_{FG}(k_1,k_2) &= \frac{F(k_1,k_2)\overline{G(k_1,k_2)}}{\left|F(k_1,k_2)\overline{G(k_1,k_2)}\right|} \\ &= e^{j\theta(k_1,k_2)}, \qquad (3) \end{aligned}$$

where $\overline{G(k_1, k_2)}$ is the complex conjugate of $G(k_1, k_2)$ and $\theta(k_1, k_2)$ denote the phase difference $\theta_F(k_1, k_2) - \theta_G(k_1, k_2)$. The POC function $r_{fg}(n_1, n_2)$ is the 2D in-

PR0001/12/0000- 0307¥400 ©2012 SICE

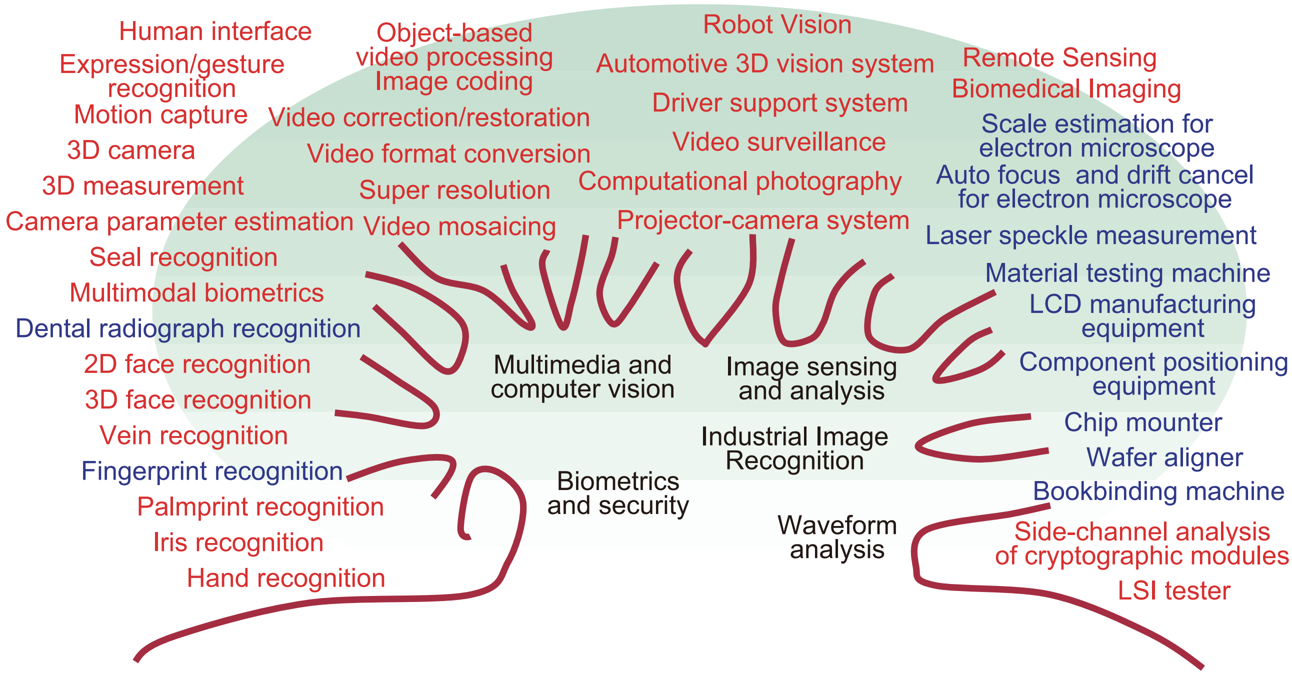

Fig. 1 Applications of POC technique.

verse DFT (IDFT) of $R_{FG}(k_1, k_2)$ and is given by

$$r_{fg}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1,k_2} R_{FG}(k_1, k_2) \times W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}, \quad (4)$$

where $\sum_{k_1,k_2}$ denotes $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2}$. When two images are similar, their POC function gives a distinct sharp peak. When two images are not similar, the peak drops significantly. The height of the peak gives a good similarity measure for image matching, and the location of the peak shows the translational displacement between the images. The key features of the POC-based image matching are summarized as follows:

(A) Translational displacement estimation

A high-accuracy translational displacement estimation method employs (i) an analytical function fitting technique to estimate the sub-pixel position of the correlation peak, (ii) a windowing technique to eliminate the effect of periodicity in DFT, and (iii) a spectrum weighting technique to reduce the effect of aliasing and noise [7]. Our experimental observation shows that POC-based matching can estimate displacement between two images with 0.01-pixel accuracy when image size is about 100 × 100 pixels.

(B) Sub-pixel correspondence search

An efficient method of sub-pixel correspondence matching employs (i) a coarse-to-fine strategy using image pyramids for robust correspondence search (Fig. 2) and (ii) a sub-pixel block matching technique for finding a pair of corresponding points with sub-pixel displacement accuracy [6, 8]. Experimental evaluation shows that the displacement between corresponding points can be estimated with 0.05-pixel accuracy when using 32 × 32-pixel matching window.

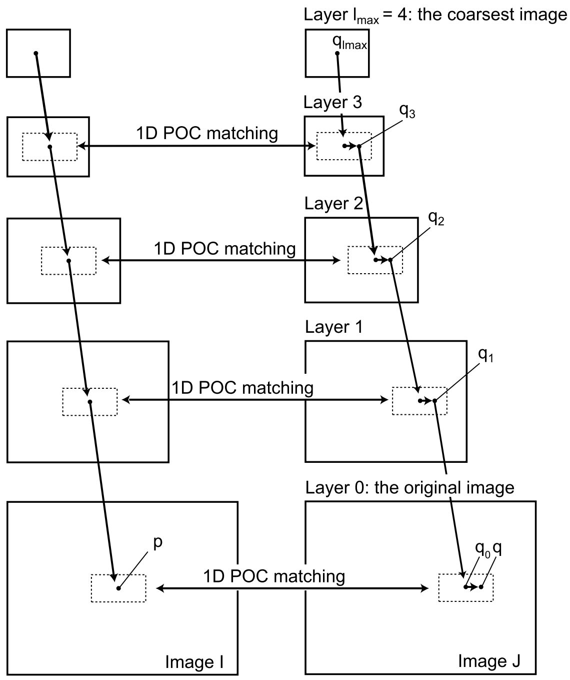

Fig. 2 A coarse-to-fine strategy using image pyramids for robust correspondence search.

(C) Rotation angle estimation

The POC-based image matching is extended to the registration of images including translation, rotation and scaling simultaneously [7]. For rotation alignment, we employ polar mapping of the amplitude spectrum to transform the image rotation into image translation. Then, the angle is estimated by detecting the translational displacement of polar mapped amplitude spectra using POC. Experimental evaluation shows the rotation angle

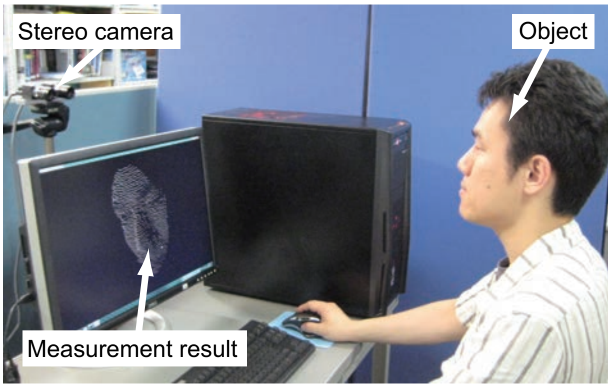

Fig. 3 Real-time 3D measurement system.

between two images can be estimated with 0.05-degree accuracy when image size is 128×128 pixels.

(D) Similarity evaluation

A similarity evaluation technique employs Band-Limited Phase-Only Correlation (BLPOC) function, which extracts the inherent frequency components of the given images and achieves highly accurate similarity evaluation for the given images [9]. Our experimental observation shows that the BLPOC function provides better discrimination capability than that of the original POC function.

3. APPLICATIONS

We present examples of potential applications of POC such as (i) real-time 3D measurement system, (ii) easy-to-use 3D measurement system using a digital camera and (iii) palmprint authentication system for mobile phones.

3.1 Real-time 3D measurement system

We have developed a real-time passive 3D measurement system using a narrow-baseline stereo camera as shown in Fig. 3. The sub-pixel correspondence search using POC allows us to achieve highly accurate 3D measurement, whose accuracy is comparable with that of laser-based active 3D measurement systems. In order to achieve real-time 3D measurement, we propose a GPU (Graphics Processing Unit) implementation of the POC-based correspondence search [12].

GPU is a specialized microprocessor that offloads and accelerates graphics rendering from the central processor (CPU). For manipulating and displaying high-quality computer graphics with a high frame rate, GPU has highly parallel structure which consists of hundreds of processors and large memory bandwidth. Recently, the highly parallel structure of GPU makes it more effective than general-purpose CPUs for algorithms where processing of large blocks of data can be done in parallel. This effort is known as General-Purpose computation on Graphics Processing Unit (GPGPU). The GPGPU has been applied to scientific computing and video processing.

We have successfully applied a GPU implementation to POC-based correspondence search due to its data parallelism and also proposed implemenation techniques to obtain the optimal performance. Through a set of experiments, the proposed GPU implementation has achieved about 4 times speed-up compared with the naive GPU implementation and about 20 times speed-up compared with the CPU implementation. Furthermore, we have demonstrated that the GPU implementation of POC-based correspondence search is more efficient than the CPU implementation by evaluating the power-deley product of each implementation.

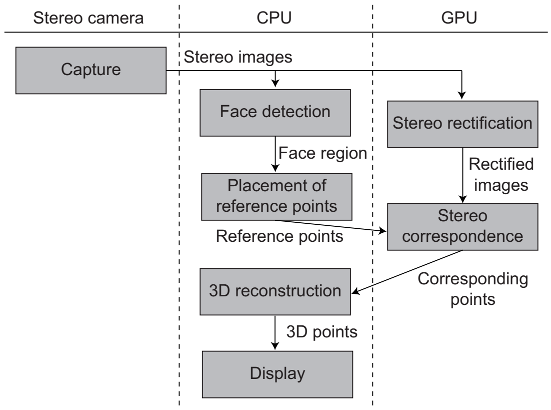

Fig. 4 Processing flow of real-time 3D face measurement system.

We have developed a real-time 3D face measurement system as shown in Fig. 4. The step of "Face detection" is to extract the face region from the captured image. The step of "Placement of reference points" is to set the reference points on the extracted face region in a reticular pattern with a spacing of 5 pixels. The camera parameters for the stereo rectification and the 3D reconstruction are obtained by the camera calibration in advance. Fig. 5 shows examples of 3D measurement results. As a result, the accurate 3D shape of the face is measured. Also, the system can measure the 3D face in real-time even if the facial expression is changed. The system can measure about 10,000 3D points in 15 fps, which is the frame rate of the camera. As is observed in the above results, the GPU implementation of the POC-based stereo correspondence matching makes it possible to achieve real-time accurate 3D measurement even with the general-purpose computer.

3.2 Easy-to-use 3D measurement system using a digital camera

We have developed an easy-to-use 3D shape measurement system using a moving and uncalibrated consumer digital camera [11]. The passive 3D measurement systems which employ binocular or multiple cameras need to select the settings of the stereo cameras depending on the target object and to calibrate cameras in advance. Hence, it is hard for the users without the technical knowledge to introduce the 3D measurement system for practical use in their daily life. On the other hand, Structure from Motion (SfM) with a moving monocular camera is known as a simple approach to measure the 3D shape of an ob-

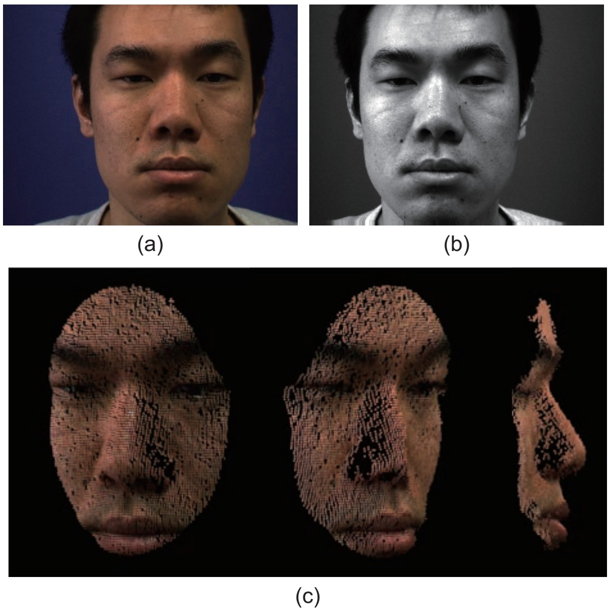

Fig. 5 3D face measurement: (a) left camera image, (b) right camera image and (c) measurement result.

ject [22, 23]. Using this approach, the user can measure 3D points of the object from multiple images captured by a monocular camera without camera calibration. However, only a limited number of 3D points are measured and are not sufficient to measure the fine 3D shape of the object, since SfM employs a feature-based correspondence matching such as Scale Invariant Feature Transform (SIFT) [24]. Addressing this problem, the developed 3D shape measurement system employs SfM using SIFT to estimate camera parameters and POC-based correspondence search to obtain dense correspondence. The use of the developed system makes it possible to measure the fine and accurate 3D shape of the object with very simple operation such as two shots of a consumer camera.

The developed system consists of a consumer digital camera and a laptop computer as shown in Fig. 6. The digital camera and the computer are connected by a Eye-Fi card [25], which is a memory card with wireless connection. First, a user captures two images of the target object with a digital camera. Next, the Eye-Fi card automatically transfers the images from the digital camera to the computer. Finally, the computer calculates 3D shape of the object from the two images. Calculating the 3D shape contains 3 steps: (i) camera parameter estimation using SIFT-based SfM, (ii) stereo rectification [22] and (iii) POC-based correspondence matching as shown in Fig. 7.

Fig. 8 shows 3D measurement results of the developed system. As a result, the developed system can measure dense 3D points compared with SIFT-based SfM. All the processes of the system is finished in about 20 seconds with a mid-range laptop computer. We have also evaluated the measurement accuracy by comparing the measurement result with the ground truth 3D model measured

Fig. 6 Easy-to-use 3D measurement system.

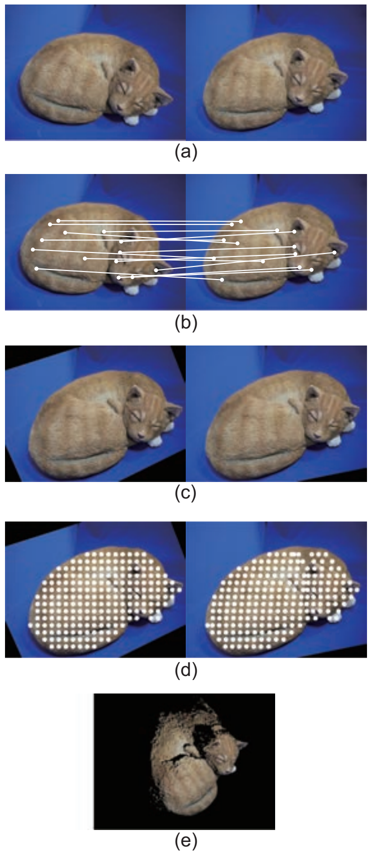

Fig. 7 Processing flow of the proposed system: (a) input stereo images, (b) result of feature-based correspondence matching, (c) rectified stereo images, (d) result of area-based correspondence matching and (e) 3D measurement result.

by the laser scanner (KONICA MINOLTA VIVID). The RMS (Root Mean Square) of the measurement error is below 1 mm.

3.3 Palmprint authentication system for mobile phones

We have developed a palmprint authentication system for mobile phones. In this system, a user takes his/her own palm image using a built-in camera on a mobile phone as shown in Fig. 9 (a). Then, the system extracts the palmprint region which is a Region Of Interest (ROI) as shown in Fig. 9 (b) and evaluates the similarity between the input and registered ROIs.

Various palmprint matching algorithms have been pro-

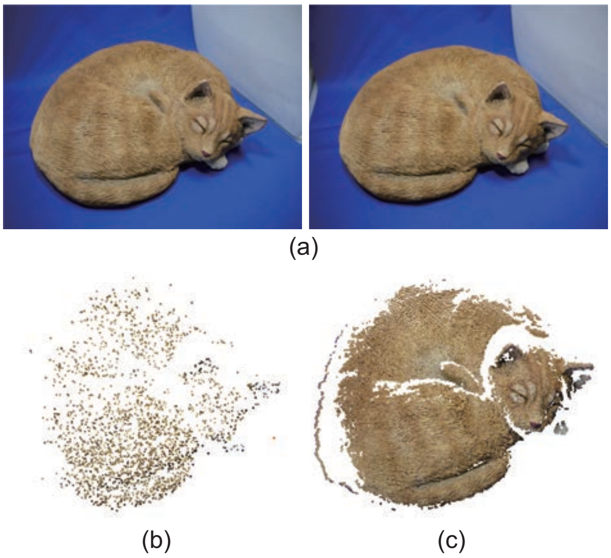

Fig. 8 3D measurement results: (a) input images, (b) conventional method (SfM using SIFT) and (c) developed system.

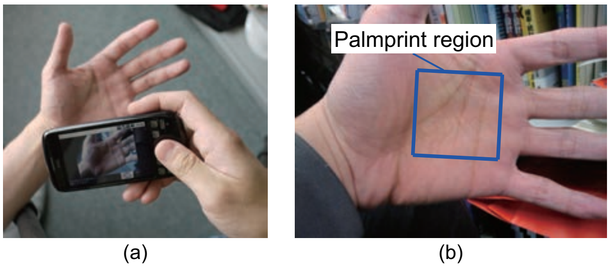

Fig. 9 Palmprint authentication system for mobile phones: (a) developed system and (b) captured image and extracted palmprint region.

posed [26]. Almost all algorithms assume that a hand is fixed on a system when a palm images are captured, and ROIs are extracted with normalizing their translation, rotation and scale variation. In practical situation such as our system, performance of conventional matching algorithms may drop significantly, since ROIs are deformed by variations in hand poses. In order to address this problem, matching algorithm of the developed system consists of 2 steps: (i) geometric correction using POC-based correspondence matching (Fig. 10 (a)) and (ii) similarity evaluation using BLPOC approximating nonlinear distortion by minute translation of local image blocks (Fig. 10 (b)).

We evaluate the performance of the matching algorithm using 520 hand images with 30 subjects and about 20 different image of each subject. The correct recognition rate of the proposed matching algorithm is about 96%, while the recognition rate under same condition of Competitive Code [27], which is one of the most successful palmprint matching algorithms, is about 92%. This result indicates that our POC-based approach is more suitable for practical situation. The total authentication time of the proposed algorithm on HTC Nexus One (CPU: Qualcomm Snapdragon QSD8250 (1GHz), RAM: 512MB, OS: Android 2.3.4, Camera: 5M resolution with autofocus) is about 1 second.

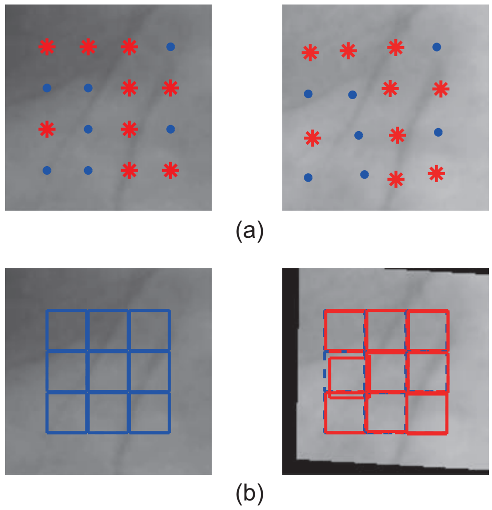

Fig. 10 Example of palmprint matching: (a) result of corresponding matching where $*$ indicates the reliable corresponding point pairs and (b) corrected ROIs and extracted local blocks for matching score calculation.

Our POC-based approach has been also successfully applied to other biometric authentication systems based on iris, fingerprint, dental radiograph, etc.

4. CONCLUSION

This paper has presented an efficient sub-pixel image matching technique based on Phase-Only Correlation (POC) and also demonstrated some examples of potential applications using POC-based image matching.

REFERENCES

[1] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," *Proc. Int'l Conf. Cybernetics and Society*, pp. 163–165, 1975.

[2] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 12, pp. 1156–1168, 1994.

[3] G. A. Thomas, "Television motion measurement for DATV and other applications," *BBC Research Department Report*, Sept. 1987.

[4] B. Kumar, A. Mahalanobis, and R. Juday, *Correlation Pattern Recognition*. Cambridge University Press, 2005.

[5] "Products using phase-based image matching." `http://www.aoki.ecei.tohoku.ac.jp/research/poc.html`.

[6] K. Takita, M. A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," *IEICE Trans. Fundamentals*, vol. E87-A, no. 8, pp. 1913–1923, 2004.

[7] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE*

Trans. Fundamentals, vol. E86-A, no. 8, pp. 1925–1934, 2003.

[8] T. Shibahara, T. Aoki, H. Nakajima, and K. Kobayashi, "A sub-pixel stereo correspondence technique based on 1d phase-only correlation," *Proc. Int'l Conf. Image Processing*, vol. 5, pp. V–221–V–224, 2007.

[9] K. Ito, H. Nakajima, K. Kobayashi, T. Aoki, and T. Higuchi, "A fingerprint matching algorithm using phase-only correlation," *IEICE Trans. Fundamentals*, vol. E87-A, no. 3, pp. 682–691, 2004.

[10] T. Takahashi, T. Kawano, K. Ito, T. Aoki, and S. Kondo, "Performance evaluation of a geometric correction method for multi-projector display using SIFT and phase-only correlation," *Proc. Int'l Conf. Image Processing*, pp. 1189–1192, 2010.

[11] S. Sakai, K. Ito, T. Aoki, and H. Unten, "Accurate and dense wide-baseline stereo matching using SW-POC," *Proc. Asian Conf. Pattern Recognition*, pp. 335–339, 2011.

[12] M. Miura, K. Fudano, K. Ito, and T. Aoki, "GPU implementation of phase-based image correspondence matching and its evaluation," *Proc. Int'l Symp. Low-Power and High-Speed Chips*, no. 19, 2011.

[13] K. Ito, A. Morita, T. Aoki, H. Nakajima, K. Kobayashi, and T. Higuchi, "Score-level fusion of phase-based and feature-based fingerprint matching algorithms," *IEICE Trans. Fundamentals*, vol. E93-A, no. 3, pp. 607–616, 2010.

[14] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An effective approach for iris recognition using phase-based image matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 30, no. 10, pp. 1741–1756, 2008.

[15] S. Iitsuka, K. Ito, and T. Aoki, "A practical palmprint recognition algorithm using phase information," *Proc. Int'l Conf. Pattern Recognition*, no. We-BCT9.18, pp. 1–4, 2008.

[16] K. Ito, T. Aoki, T. Hosoi, and K. Kobayashi, "Face recognition using phase-based correspondence matching," *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pp. 173–178, 2011.

[17] S. Aoyama, K. Ito, and T. Aoki, "Finger-knuckle-print recognition using BLPOC-based local block matching," *Proc. Asian Conf. Pattern Recognition*, pp. 525–529, 2011.

[18] K. Ito, Y. Hanzawa, T. Aoki, M. Inagaki, R. Kawamata, and E. Kosuge, "A practical human identification method using dental radiographs and dental CT data," *Proc. Radiological Society of North America*, p. 1069, 2009.

[19] K. Miyazawa, Y. Tajima, K. Ito, T. Aoki, A. Katsumata, and K. Kobayashi, "A novel approach for volume registration using 3D phase-only correlation," *Proc. Radiological Society of North America*, p. 1070, 2009.

[20] N. Homma, S. Nagashima, T. Sugawara, T. Aoki, and A. Satoh, "A high-resolution phase-based waveform matching and its application to side-channel attacks," *IEICE Trans. Fundamentals*, vol. E91-A, no. 1, pp. 193–202, 2008.

[21] N. Homma, A. Miyamoto, T. Aoki, A. Satoh, and A. Shamir, "Comparative power analysis of modular exponentiation algorithms," *IEEE Trans. Comput.*, vol. 59, no. 6, pp. 795–807, 2010.

[22] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, 2010.

[23] R. Hartley, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2008.

[24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int'l J. Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[25] Eye-Fi Inc., "Eye-Fi memory cards." `http://www.eye.fi`.

[26] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition*, vol. 42, no. 7, pp. 1408–1418, 2009.

[27] A. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," *Proc. Int'l Conf. Pattern Recognition*, vol. 1, pp. 520–523, 2004.